HOW TO STAY IN A SET

OR

KONIG'S LEMMA FOR RANDOM PATHS

by

Roger A. Purves and William D. Sudderth*

University of Minnesota
Technical Report 375

June 1980

* Research supported by National Science Foundation Grant MCS 77-28424 .

ABSTRACT

Starting at state $x \in X$, a player selects the next state x_1 from the collection $\Gamma(x)$ of those available and then selects x_2 from $\Gamma(x_1)$ and so on. Suppose the object is to control the path $x_1, x_2, \ldots$ so that every x_i will lie in a subset A of X. A famous lemma of König is equivalent to the statement that if every $\Gamma(x)$ is finite and if, for every n, the player can obtain a path in A of length n, then the player can obtain an infinite path in A. Here paths are not necessarily deterministic and, for each x, $\Gamma(x)$ is the collection of possible probability distributions for the next state. Under mild measurability conditions, it is shown that if, for every n, there is a random path of length n which lies in A with probability larger than α, then there is an infinite random path with the same property. Furthermore, the measurability and finiteness assumptions can be dropped if, in the hypothesis, the positive integers n are replaced by stop rules t. An analogous result holds when the object is to visit A infinitely many times.

Key words: probability, gambling, dynamic programming, stochastic control, finite additivity, regularity of set functions, game theory.

1. Introduction and statement of results. Suppose X is a nonempty set of possible states for a process and that to each $x \in X$ is associated a nonempty collection $\Gamma(x)$ of finitely additive probability measures defined on all subsets of X. Then, starting from any x, one can construct a random sequence $x_1, x_2, x_3, \cdots$ by selecting $\sigma_0 \in \Gamma(x)$ to be the distribution of x_1, then selecting $\sigma_1(x_1) \in \Gamma(x_1)$ to be the conditional distribution of x_2 given x_1, and $\sigma_2(x_1, x_2) \in \Gamma(x_2)$ to be the conditional distribution of x_3 given x_1, x_2, and so on. The sequence $\sigma = \{\sigma_0, \sigma_1, \cdots\}$ is a strategy at x in Γ. As is explained in Dubins and Savage (1976 , pages 7-21), Dubins (1974), and Purves and Sudderth (1976), each strategy determines a finitely additive probability measure on the sigma-field $\mathcal{B}$ of subsets of $H = X \times X \times \cdots$ which is generated by the open subsets of H when X is assigned the discrete topology and H the product topology. This measure is also denoted by σ and is regarded as the distribution of the random sequence $x_1, x_2, \cdots$. (A reader unfamiliar with finite additivity theory can assume X countable and all measures countably additive on all subsets of X to get the gist of the results.)

Let g be a bounded, real-valued $\mathcal{B}$-measurable function defined on H and, for $h = (h_1, h_2, \cdots) \in H$, regard $g(h)$ as the payoff to the player from the sequence h. The optimal reward operator is defined to be that functional Γ^∞ which assigns to the function g on H the function $\Gamma^\infty g$ on X defined by

$$(\Gamma^\infty g)(x) = \sup\{\sigma g: \sigma \text{ at } x\}$$

for each $x \in X$. Thus $(\Gamma^\infty g)(x)$ is the supremum of the possible expected payoffs for a player starting at x.

Throughout this paper we will use de Finetti's convention of identifying a set B with its indicator function. So, for example, $\Gamma^\infty B$ is written instead of $\Gamma^\infty 1_B$ for a set $B \in \mathcal{B}$.

Our first result relies on countable additivity. To state it, assume that $\mathcal{A}$ is a sigma-field of subsets of X and identify each measure defined on all subsets of X with its restriction to $\mathcal{A}$. Suppose that, for each $n = 1,2,\cdots$, γ_n is a mapping which assigns to each $x \in X$ a countably additive probability measure $\gamma_n(x)$ on $\mathcal{A}$. Assume that each γ_n is $\mathcal{A}$-measurable in the sense that $\gamma_n(x)(A)$ is an $\mathcal{A}$-measurable function of x for every $A \in \mathcal{A}$. Further assume that, for every x,

$$\Gamma(x) = \{\gamma_n(x): n = 1,2,\cdots\}$$

and that there is an $n = n(x)$ such that $\gamma_k(x) = \gamma_n(x)$ for all $k \geq n$. Such a Γ is said to be $\mathcal{A}$-measurable and pointwise finite.

For any subset A of X and for $n = 1,2,\cdots$, let

$$A^n = \{h \in H: h_i \in A, \ i = 1,2,\cdots,n\}$$

$$A^\infty = \{h \in H: h_i \in A, \ i = 1,2,\cdots\} \quad .$$

Theorem 1. If Γ is $\mathcal{A}$-measurable and pointwise finite and if $A \in \mathcal{A}$, then

$$\Gamma^\infty A^\infty = \inf_n \Gamma^\infty A^n \quad . \tag{1.1}$$

To see that Theorem 1 is a generalization of König's lemma

(König (1936)), suppose that all probability measures available are pointmasses and take $\mathcal{A}$ to be the sigma-field of all subsets of X. (Interesting modern treatments of König's lemma are in Knuth (1973) and Kuratowski and Mostowski (1976).) Theorem 1 is almost a corollary to results of Schäl (1975) and Schreve and Bertsekas (1979) where more general payoff functions are considered. However, we allow for a more general sigma-field and our proof is somewhat simpler.

As is well-known and easy to see, König's Lemma may fail if the condition of finiteness is dropped. In addition, (1.1) may also fail if the assumption of countable additivity is not satisfied. However, an analogous result holds in a general setting if the integers are replaced by stop rules.

An <u>incomplete stop rule</u> τ, as defined in [3], is a mapping $\tau: H \to \{1,2,\cdots\} \cup \{\infty\}$ such that if $\tau(h) < \infty$ and $h_i = h'_i$, $i = 1,\cdots,\tau(h)$, then $\tau(h) = \tau(h')$. A <u>stop rule</u> t is an incomplete stop rule which is everywhere finite. If t is a stop rule and $A \subset X$, let

$$A^t = \{h: h \in A^{t(h)}\} .$$

<u>Theorem 2.</u> For every Γ and every $A \subset X$,

$$\Gamma^\infty A^\infty = \inf_t \Gamma^\infty A^t . \tag{1.2}$$

Furthermore, for each $\varepsilon > 0$, there is a stop rule t such that $(\Gamma^\infty A^t)(x) \le (\Gamma^\infty A^\infty)(x) + \varepsilon$ for all $x \in X$.

Equation (1.2) says that if, for each stop rule t, there is a way to stay in A at least until time t with probability at least 1/2 (say), then there is a way to stay in A forever with probability at least $1/2 - \varepsilon$.

To state our final result, let $A \subset X$ and define

$$[A \text{ i.o.}] = \{h: h_i \in A \text{ for infinitely many } i\} \quad .$$

Theorem 3. For every Γ and every $A \subset X$,

$$(1.3) \qquad \Gamma^{\infty}[A \text{ i.o.}] = \inf_{\tau} \Gamma^{\infty} [\tau < \infty]$$

where the infimum is over all incomplete stop rules τ which are finite on every h in $[A \text{ i.o.}]$. Furthermore, for each $\varepsilon > 0$, there is such a τ for which

$$\Gamma^{\infty}[\tau < \infty](x) \leq \Gamma^{\infty}[A \text{ i.o.}](x) + \varepsilon$$

for all $x \in X$.

Each of the three theorems can be regarded as approximation results in the Dubins and Savage theory of gambling. In particular, the left-hand-side of (1.3) is the optimal return function, denoted by V in Dubins and Savage (1976), for the nonleavable gambling problem which has gambling house Γ and utility function A.

The next section has some preliminary definitions and results about the optimal reward operator. The succeeding three sections present proofs of the theorems. The final one has additional remarks and some open questions.

2. Two properties of the optimal reward operator. Throughout this section g, g_1, and g_2 are bounded, real-valued, $\mathcal{B}$-measurable functions with domain H. Here is a trivial but useful property of Γ^∞.

Lemma 1. $\qquad g_1 \leq g_2 \Rightarrow \Gamma^\infty g_1 \leq \Gamma^\infty g_2$.

If $p = (x_1,\cdots,x_m)$, then gp is the function on H defined by $(gp)(h) = g(ph)$ where ph is that element of H which consists of the terms of p followed by those of h. For a stop rule t and $h = (h_1,h_2,\cdots) \in H$, $h_t(h) = h_{t(h)}$ and $p_t(h) = (h_1,\cdots,h_t(h))$. For a strategy σ, $\sigma[p]$ denotes the conditional strategy given p and $\sigma(g|p)$ is often written for the quantity $\sigma[p](gp)$, which it is natural to regard as the conditional σ-expectation of g. Thus the formula

$$(2.1) \qquad \sigma g = \int \sigma(g|p_t)d\sigma$$

can be interpreted in the usual way as conditioning on the past up to time t. It was proved for finitary (continuous) g in [3] (equation 3.7.1) and holds for all bounded, $\mathcal{B}$-measurable g as follows from Theorems 4.1 and 5.1 of [9]. The next lemma gives a similar formula for the operator Γ^∞ and is a version of the optimality equation of dynamic programming. To state it, define $\Gamma^\infty(g|p_t)$ to be that function on H whose value at h is $(\Gamma^\infty gp)(x)$ when $p = p_t(h)$ and $x = h_t(h)$. Notice that $\Gamma^\infty(g|p_t)$ is determined by time t and, hence is a finitary function (Theorem 2.7.1, [3]). It is natural to regard $\Gamma^\infty(g|p_t)(h)$ as the conditional optimal reward which it is possible for a player to achieve who has so far experienced $p_t(h)$.

For the proof of the lemma, another definition is needed. Two strategies σ and σ' <u>agree prior to a time</u> t if $\sigma_0 = \sigma_0'$ and, for every h and n with $0 < n < t(h)$, $\sigma_n(p_n(h)) = \sigma_n'(p_n(h))$.

<u>Lemma 2.</u> $\Gamma^\infty g = \Gamma^\infty(\Gamma^\infty(g|p_t))$.

Pf: For any σ at x,

$$\begin{aligned}\sigma g &= \int \sigma(g|p_t)d\sigma \\ &\leq \int \Gamma^\infty(g|p_t)d\sigma \\ &\leq \Gamma^\infty(\Gamma^\infty(g|p_t)) \quad .\end{aligned}$$

Take the supremum over σ at x to see that $(\Gamma^\infty g)(x) \leq \Gamma^\infty(\Gamma^\infty(g|p_t))(x)$. To prove the opposite inequality, let $\varepsilon > 0$. Let σ be a strategy at x such that

$$\sigma(\Gamma^\infty(g|p_t)) \geq \Gamma^\infty(\Gamma^\infty(g|p_t))(x) - \varepsilon/2$$

and, for each $h \in H$, let $\bar{\sigma}(h) = \bar{\sigma}(p_t(h))$ be a strategy at $h_t(h)$ such that

$$\bar{\sigma}(h)(gp_t(h)) \geq \Gamma^\infty(g|p_t)(h) - \varepsilon/2.$$

Define σ' to be that strategy at x which agrees with σ prior to time t and has the conditional strategy $\sigma'[p_t(h)] = \bar{\sigma}(h)$ for every h. Then

$$\begin{aligned}(\Gamma^\infty g)(x) &\geq \sigma' g \\ &= \int \sigma'(g|p_t)d\sigma'\end{aligned}$$

$$= \int \bar{\sigma}(h)(gp_t(h))d\sigma$$

$$\geq \int \Gamma^{\infty}(g|p_t)d\sigma - \varepsilon/2$$

$$\geq \Gamma^{\infty}(\Gamma^{\infty}(g|p_t)) - \varepsilon. \quad \square$$

A counterpart of Lemma 2 for Borel measurable problems is Lemma 4.6 of Dubins and Sudderth (1977).

For a bounded, real-valued function φ defined on X, let

$$(\Gamma^1\varphi)(x) = \sup\{\gamma\varphi: \gamma \in \Gamma(x)\} \quad .$$

If t is the constant 1, then the function $\Gamma^{\infty}(g|p_t) = \Gamma^{\infty}(g|p_1)$ depends only on the first coordinate h_1 of h and can be regarded as a function defined on X. With this proviso, the equality of Lemma 2 specializes to give

$$\Gamma^{\infty}g = \Gamma^1(\Gamma^{\infty}(g|p_1)) \quad . \tag{2.2}$$

3. The proof of Theorem 1. In this section Γ is assumed to be A-measurable and pointwise finite so that $\Gamma(x) = \{\gamma_n(x) \colon n = 1,2,\cdots\}$ where the γ_n are A-measurable and $\gamma_k(x) = \gamma_n(x)$ for $k \geq n = n(x)$. Also $A \in A$.

Let $Q = \inf_n \Gamma^\infty A^n$.

Lemma 1. $\Gamma^\infty A^\infty \leq Q$.

Proof: For every n, $A^\infty \subseteq A^n$ and, hence, $\Gamma^\infty A^\infty \leq \Gamma^\infty A^n$. □

The proof of the opposite inequality will be given in several lemmas.

Lemma 2. For every x, $(\Gamma^\infty A^1)(x) = (\Gamma^1 A)(x)$ and, for $n \geq 1$,

$$(\Gamma^\infty A^{n+1})(x) = \Gamma^1(A(\Gamma^\infty A^n))(x).$$

Proof: Use (2.2) □

Lemma 3. For every n, $\Gamma^\infty A^n$ is A-measurable and, hence, Q is also.

Proof: If g is bounded, A-measurable, then $x \to \gamma_k(x)g$ is A-measurable, and, consequently, $x \to (\Gamma^1 g)(x) = \sup \gamma_k(x)g$ is A-measurable. Now use Lemma 2. □

Lemma 4. If $\{g_i\}$ is a uniformly bounded sequence of A-measurable functions defined on X and converging pointwise to g, then $\lim_i(\Gamma^1 g_i)(x) = (\Gamma^1 g)(x)$ for all x.

Proof: Fix x and $n = n(x)$. Then

$$\lim_i (\Gamma^1 g_i)(x) = \lim_i \max_{k \le n} \gamma_k(x) g_i$$

$$= \max_{k \le n} \lim_i \gamma_k(x) g_i$$

$$= \max_{k \le n} \gamma_k(x) g$$

$$= (\Gamma^1 g)(x). \qquad \square$$

<u>Lemma 5.</u> $Q = \Gamma^1(AQ)$.

Proof: Let $n \to \infty$ in the second equation of Lemma 2 and make use of Lemmas 3 and 4. $\square$

By Lemma 5 and the pointwise finiteness of Γ, there is, for every x, a $\gamma \in \Gamma(x)$ such that

(3.1) $\gamma(AQ) = Q(x)$.

Let $k(x)$ be the least k such that (3.1) holds with $\gamma = \gamma_k(x)$ and let $\gamma(x) = \gamma_{k(x)}(x)$ for every x. For each x, let $\sigma = \bar{\sigma}(x)$ be the strategy which uses $\gamma(y)$ whenever the current state is y.

<u>Lemma 6.</u> For every x, $\bar{\sigma}(x)(A^\infty) \ge Q(x)$.

Proof: Use (3.1) and induction to see that, for every n,

$$\int A^n(h) Q(h_n) d\bar{\sigma}(x)(h) = Q(x) .$$

Since $0 \le Q \le 1$,

$$\bar{\sigma}(x)(A^n) \ge Q(x)$$

for all n. Now let $n \to \infty$ and use the fact that $\bar{\sigma}(x)$ is countably additive on A^{∞}. $\square$

Because $\bar{\sigma}(x)$ is a strategy at x,

$$(\Gamma^{\infty} A^{\infty})(x) \geq \bar{\sigma}(x)(A^{\infty})$$

and Theorem 1 follows from Lemmas 1 and 6 Notice that it has also been shown that $\bar{\sigma}$ is an optimal stationary family of strategies. The existence of such a family follows more directly from Theorem 3.9.6 of Dubins and Savage (1976).

4. <u>The proof of Theorem 2.</u> In this section, let

$$Q = \inf\{\Gamma^\infty A^t : t \text{ a stop rule}\}$$

<u>Lemma 1.</u> $\Gamma^\infty A^\infty \leq Q$.

Proof: $A^\infty \subset A^t$ for every t. □

<u>Lemma 2.</u> $Q \leq \Gamma^1(AQ)$. Indeed, for every $\varepsilon > 0$, there is a t such that $(\Gamma^\infty A^t)(x) \leq \Gamma^1(AQ)(x) + \varepsilon$ for all x.

Proof: Fix $\varepsilon > 0$. By definition of Q, there is, for every x, a stop rule $\bar{t}(x)$ such that

$$(\Gamma^\infty A^{\bar{t}(x)})(x) \leq Q(x) + \varepsilon.$$

Define the stop rule t by

$$t(h) = \bar{t}(h_1)(h_2, h_3, \cdots) + 1.$$

Then

$$\begin{aligned}(A^t)h_1 &= A^{\bar{t}(h_1)} \quad \text{if } h_1 \in A\,, \\ &= \phi \quad\quad\ \text{if } h_1 \notin A\,.\end{aligned}$$

So, by (2.2),

$$\begin{aligned}\Gamma^\infty A^t &= \Gamma^1(\Gamma^\infty(A^t|h_1)) \\ &= \Gamma^1(A(\Gamma^\infty A^{\bar{t}(h_1)})) \\ &\leq \Gamma^1(AQ) + \varepsilon\,.\end{aligned}$$ □

Now let $\varepsilon > 0$. Use Lemma 2 to find for every x and $n = 0,1,\cdots$

an element $\gamma_n(x) \in \Gamma(x)$ such that $\gamma_n(x)(AQ) \geq Q(x) - \varepsilon/2^{n+1}$. Next for each x, let $\sigma = \bar{\sigma}(x)$ be the strategy at x which has $\sigma_0 = \gamma_0(x)$ and $\sigma_n(x_1,\ldots,x_n) = \gamma_n(x_n)$ for every n and $x_1,\ldots,x_n$.

Lemma 3. For every stop rule t and every $x \in X$,

$$Q(x) \leq \bar{\sigma}(x)(A^t Q(h_t)) + \varepsilon \leq \bar{\sigma}(x)(A^t) + \varepsilon \quad .$$

Proof: The sequence $Q_n(h)$ defined by $Q_0 = Q(x)$ and $Q_n(h) = A^n(h)Q(h_n) + \varepsilon(1 - 1/2^n)$ is an upper semimartingale under $\bar{\sigma}(x)$. The first inequality now now follows from Theorem 2.12.2 of [3]. The second is obvious. □

By corollary 5.3 of [9], $\bar{\sigma}(x)(A^t)$ converges to $\bar{\sigma}(x)(A^\infty)$ as t increases. Also, $\bar{\sigma}(x)(A^\infty) \leq (\Gamma^\infty A^\infty)(x)$ by definition of Γ^∞. Thus it follows from Lemma 3 that

$$Q \leq \Gamma^\infty A^\infty \ .$$

This together with Lemma 1 yields the first assertion of Theorem 2. For the second assertion, first use (2.2) to see that

$$\Gamma^\infty A^\infty = \Gamma^1(A(\Gamma^\infty A^\infty)) = \Gamma^1(AQ)$$

and then use the second assertion of Lemma 2.

5. <u>The proof of Theorem 3.</u> Let $A \subset X$ and let

$$G^{\infty} = [A \text{ i.o.}] \ .$$

Also, in this section, let

$$Q = \inf\{\Gamma^{\infty}[\tau < \infty]: \tau \in T\}$$

where T is the collection of incomplete stop rules τ such that $G \subset [\tau < \infty]$. By Lemma 2.1,

$$\Gamma^{\infty}G^{\infty} \leq Q \ . \tag{5.1}$$

Most of the remainder of this section is devoted to proving the opposite inequality.

Here is a simple technical lemma.

<u>Lemma 1.</u> If τ is an incomplete stop rule and σ is a strategy, then

$$\begin{aligned} \sigma[\tau < \infty] &= \sup\{\sigma[\tau \leq s]: s \text{ a stop rule}\} \\ &= \sup\{\sigma[\tau = r]: r \text{ a stop rule}\} \ . \end{aligned} \tag{5.2}$$

If, in addition, $B \in \mathcal{B}$ is a subset of $[\tau < \infty]$, then

$$\begin{aligned} \sigma(B) &= \int_{\tau<\infty} \sigma(B|p_{\tau})d\sigma \\ &= \sup\{\int_{\tau=r} \sigma(B|p_{\tau})d\sigma: r \text{ a stop rule}\} \ . \end{aligned} \tag{5.3}$$

Proof: The sets $[\tau \leq n]$ increase to $[\tau < \infty]$, and so, by Corollary 5.3 of [9], $\sigma[\tau \leq s]$ increases to $\sigma[\tau < \infty]$. This establishes the first equality in (5.2). The second follows from the fact that $[\tau = r] = [\tau \leq s]$

when r is taken to be $\tau\wedge s$, the minimum of τ with x.

To see (5.2), let s be a stop rule, $r = \tau\wedge s$, and use (2.1) to write

$$\sigma(B) = \int\sigma(B|p_{\tau\wedge s})d\sigma$$

$$= \int_{\tau\leq s}\sigma(B|p_{\tau})d\sigma + \int_{\tau>s}\sigma(B|p_{s})d\sigma \quad .$$

The second term on the right equals

$$\sigma(B\cap[\tau > s]) \leq \sigma[s < \tau < \infty]$$

and approaches 0 as s increases by (5.2). The first term approaches

$$\int_{\tau<\infty}\sigma(B|p_{\tau})d\sigma \quad . \qquad \square$$

The next lemma is crucial to the proof of the inequality opposite to (5.1). Let t be the incomplete stop rule corresponding to the time of first entrance into A and let

$$G^1 = [t < \infty] \ .$$

So G^1 is the event that at least one visit is made to A.

<u>Lemma 2.</u> $Q \leq \Gamma^{\infty}(G^1(Q(h_t)))$. Indeed, for each $\varepsilon > 0$, there is a $\tau \in T$ such that

$$(\Gamma^{\infty}[\tau > \infty])(x) \leq (\Gamma^{\infty}(G^1Q(h_t)))(x) + \varepsilon$$

for all x.

Proof: Let $\varepsilon > 0$. For each $x \in A$, choose a $\bar{\tau}(x) \in T$ such that

$$(\Gamma^{\infty}[\bar{\tau}(x) < \infty])(x) < Q(x) + \varepsilon/2 \qquad .$$

Define

$$\tau(h) = \infty \text{ if } t(h) = \infty$$

$$= t(h) + \bar{\tau}(h_t(h))(h_{t(h)+1}, h_{t(h)+2}, \ldots)$$
$$\text{if } t(h) < \infty \quad .$$

Then $\tau \in T$ and, for each x, there is, by definition of Q, a σ at x such that

(5.3) $$Q(x) \leq \sigma[\tau < \infty] + \varepsilon/2.$$

Now use (5.2) to calculate

(5.4) $$\sigma[\tau < \infty] = \int_{t<\infty} \sigma(\tau < \infty | p_t) d\sigma$$
$$= \int_{t<\infty} \sigma[p_t][\bar{\tau}(h_t) < \infty] d\sigma$$
$$\leq \int_{t<\infty} Q(h_t) d\sigma + \varepsilon/2 \ .$$

The desired result is a consequence of (5.3) and (5.4). □

In the sequel it will be shown that any function Q which has values in [0,1] and satisfies the inequalities of Lemma 2 is dominated by $\Gamma^\infty G^\infty$.

The next lemma restates part of Lemma 2 in a more useful form.

Lemma 3. For every $x \in X$ and $\varepsilon > 0$, there is a $\sigma = \bar{\sigma}(x,\varepsilon)$ at x and a stop rule $r = \bar{r}(x,\varepsilon)$ such that

(5.5) $$\int_K Q(h_r) d\sigma > Q(x) - \varepsilon$$

where $K = K(x,\varepsilon) = [t = r]$ is a clopen subset of $[h_r \in A]$ and is determined by time r.

Proof: By Lemma 2, there is a σ at x such that

$$\int_{t<\infty} Q(h_t)d\sigma > Q(x) - \varepsilon/2 .$$

Now use (5.2) to approximate $[t < \infty]$ by $[t = r]$. □

Let $\varepsilon > 0$ and $x_0 \in X$. The result of Lemma 3 will now be used to construct a strategy σ at x such that

(5.6) $$\sigma(G^\infty) \geq Q(x) - \varepsilon.$$

This inequality is enough to establish the reverse inequality to (5.1)

The definition of σ depends on an inductive construction of a sequence $\{K_n\}$ of clopen sets and a sequence $\{r_n\}$ of stop rules which have the following properties.

(i) For every n, K_n is determined by time r_n.

(ii) For every h and n, $r_n(h) < r_{n+1}(h)$.

(iii) $\int_{K_1} Q(h_{r_1})d\sigma > Q(x) - \varepsilon/2$

(iv) For $n = 1,2,\cdots$ and $h \in \bigcap_1^n K_i$,

$$\int_{K_{n+1}q_n(h)} Q(h_{r_{n+1}[q_n(h)]})d\sigma[q_n(h)] > Q(h_{r_n}(h)) - \varepsilon/2^n ,$$

where $q_n(h)$ is written for $p_{r_n}(h)$.

(v) $\bigcap_1^\infty K_i \subset G^\infty$.

With the aid of Lemma 4, it is not difficult to carry out the construction.

Take $r_1 = \bar{r}(x, \varepsilon/2)$, $K_1 = K(x, \varepsilon/2)$ and let σ agree with $\bar{\sigma}(x, \varepsilon/2)$ prior to time r_1. By Lemma 3,

$$(5.7) \qquad K_1 \subset [h_{r_1} \in A] .$$

Suppose $r_1,\ldots,r_n$ and $K_1,\ldots,K_n$ have been defined. Take r_{n+1} to be that stop rule which, given the past up to r_n, continues to time $\bar{r}(h_{r_n}, \varepsilon/2^n)$. That is,

$$r_{n+1}[q_n(h)] = \bar{r}(h_{r_n}(h), \varepsilon/2^n).$$

for all h. Also take K_{n+1} to be that clopen set which satisfies

$$K_{n+1}q_n(h) = K(h_{r_n}(h), \varepsilon/2^n) .$$

Notice that

$$(5.8) \qquad K_{n+1} \subset [h_{r_{n+1}} \in A] .$$

Fonally require the conditional strategy $\sigma[q_n(h)]$ to agree with $\bar{\sigma}(h_{r_n}(h), \varepsilon/2^n)$ prior to time $\bar{r}(h_{r_n}(h), \varepsilon/2^n)$.

Properties (i) and (ii) are clear from the construction. Properties (iii) and (iv) are instances of (5.5). Property (v) is a consequence of (5.8).

The next lemma generalizes Lemma 7.1 of [9] and the proof is similar.

<u>Lemma 4.</u> Let Q be a function from X to the interval $[0,1]$ and

suppose that Q, ε, x, $\{K_n\}$, and $\{r_n\}$ have properties (i) through (iv). Then

$$\sigma(\bigcap_1^\infty K_i) \geq Q(x) - \varepsilon \quad . \tag{5.9}$$

Proof: There is no harm in assuming, as we do, that $Q(x) > \varepsilon$.

Let K be clopen and contain $\bigcap_1^\infty K_i$. It suffices to show

$$\sigma(K) \geq Q(x) - \varepsilon \quad . \tag{5.10}$$

The proof of (5.10) is by induction on the structure of K.

Suppose first that K has structure 0. Then either $K = H$ or $K = \phi$. If $K = H$, (5.10) is clear. To show that K cannot be empty, a history $h \in \cap K_i$ will be constructed. By property (iii), there is an $h^1 \in K_1$ such that $Q(h_{r_1}(h^1)) \geq Q(x) - \varepsilon/2 > \varepsilon/2$. So, by property (iv), there is $h^2 \in K^2$ such that h^2 agrees with h^1 up to time $r_1(h^1)$ (i.e. $q_1(h^1) = q_1(h^2)$) and $Q(h_{r_2}(h^2)) \geq Q(h_{r_1}(h^1)) - \varepsilon/4 > \varepsilon/4$. Continue in this fashion to define $h^{n+1} \in K_{n+1}$ such that h^{n+1} agrees with h^n up to time $r_n(h^n)$. Then take h to be that history which agrees with h^n up to time $r_n(h^n)$ for every n. Since K_n is determined by time r_n and $h^n \in K_n$, the history h is in K_n for every n.

For the inductive step, assume (5.10) holds for sets K having structure less than the positive ordinal α, and then suppose K has structure α.

Fix $h \in K_1$ for this paragraph and set $x' = h_{r_1}(h)$, $\varepsilon' = \varepsilon/2$, $q = q_1(h)$, $\sigma' = \sigma[q]$, $K'_n = K_{n+1}q$, $r'_n = r_{n+1}[q]$, and $K' = Kq$. This new collection with primes satisfies all the hypotheses of the lemma. In addition, K' contains $\bigcap_1^\infty K'_n$ and has structure less than α. So, by the inductive hypothesis,

$$(5.11) \qquad \sigma(K|q_1(h)) = \sigma'(K') \geq Q(h_{r_1}(h)) - \varepsilon/2.$$

Now use (2.1) and calculate.

$$\sigma(K) = \int\sigma(K|q_1(h))d\sigma(h)$$

$$\geq \int_{K_1} Q(h_{r_1})d\sigma - \varepsilon/2$$

$$> Q(x) - \varepsilon .$$

The first inequality is by (5.11) and the second is by property (i) . □

By Lemma 4 and property (v),

$$\sigma(G^\infty) \geq \sigma(\bigcap_1^\infty K_i) \geq Q(x) - \varepsilon .$$

Thus, because ε is arbitrary and σ is a strategy at x,

$$(5.12) \qquad (\Gamma^\infty G^\infty)(x) \geq Q(x) .$$

This inequality together with (5.1) implies

(1.3) In the notation of this section, $Q = \Gamma^\infty G^\infty$. The final assertion of Theorem 3 follows from Lemma 2 and the lemma below.

<u>Lemma 5.</u> $Q \geq \Gamma^\infty(G^1 Q(h_t))$

Proof: Let $x \in X$ and let σ be a strategy at x. It suffices to show

$$\sigma(G^1 Q(h_t)) = \int_{t<\infty} Q(h_t)d\sigma$$

$$\leq Q(x)$$

By Lemma 1, t can be approximated by a stop rule r and it suffices to show

$$\int Q(h_r)d\sigma \leq Q(x) \quad .$$

This holds by Corollary 3.3.4 of Dubins and Savage (1976) which applies because $Q = \Gamma^{\infty}G^{\infty}$ is the optimal return function V of the gambling problem with utility function A. □

6. Remarks on regularity, measurability and game theory. For a fixed x, the optimal reward operator determines a set function

$$\mu_x(B) = (\Gamma^{\infty}B)(x)$$

on the sets $B \in \mathcal{B}$. Each of the three theorems can be interpreted in terms of the regularity (and uniform regularity) of these set functions. For example, if $A \subset X$, then the set A^{∞} is closed in H and, for each stop rule t, the set A^t is clopen. Thus Theorem 2 states that the value of μ_x on the closed set A^{∞} is the infimum of its values on clopen sets containing it. Similarly, Theorem 3 states that the G_{δ} set [A i.o.] can be μ_x-approximated by an open set containing it. We believe that these results still hold when A^{∞} is replaced by an arbitrary closed set and [A i.o.] by an arbitrary G_{δ}. It would be interesting to know whether every $B \in \mathcal{B}$ can be μ_x-approximated by open sets containing it, but we do not know the answer even in the special case when X is finite.

Presumably, Theorems 2 and 3 remain true in a Borel measurable, countably additive setting where X is standard Borel, Γ is a Borel house as in Strauch (1967), A is a Borel subset of X, and only Borel measurable stopping times are considered. There are, however, measure theoretic difficulties in adapting the proofs. For example, we have not succeeded in proving the universal measurability of the functions Q in sections 4 and 5 except in the special case when Γ is Borel absolutely continuous as in Dubins and Sudderth (1979). (Additional references for measurable gambling and dynamic programming are Blackwell, Freedman, and Orkin (1974) and Dubins and Sudderth (1977).)

All three of the theorems can be interpreted as statements that certain games have values. For example, here is equality (1.3) in a different guise.

$$\sup_{\sigma} \inf_{O} \sigma(O) = \inf_{O} \sup_{\sigma} \sigma(O) \tag{6.1}$$

The supremum is over all σ at x in Γ and the infimum is over all open sets O which contain $G = [A \text{ i.o.}]$. (Every such O is of the form $[\tau < \infty]$ for some incomplete stop rule τ.) In the terminology of game theory, player A chooses an open set containing G, player B chooses a strategy at x, and A pays B an amount equal to the measure of the open set under the strategy. Equality (6.1) expresses the fact that this game has a value. However, if σ is allowed to vary over an arbitrary collection of measures rather than the strategies at some x, then a simple example shows that the game need not have a value.

<u>Example</u>. Let $X = \{0,1\}$, $A = \{0\}$, and, for every $n = 1,2,\cdots$, let λ_n be point-mass at that element $h^{(n)} \in H$ such that $h_i^{(n)} = 0$ for $i \leq n$ and $= 1$ for $i > n$. Then, for every open set $O \supset G$, there is an n such that $h^{(n)} \in O$ and, hence,

$$\inf_{O} \sup_{n} \lambda_n(O) = 1 .$$

However,

$$\sup_{n} \inf_{O} \lambda_n(O) = \sup_{n} \lambda_n(G) = 0 .$$

REFERENCES

1. Blackwell, D., Freedman, D., and Orkin, M. (1974). The optimal reward operator in dynamic programming. Ann. Probability 2 926 - 941.

2. Dubins, Lester E. (1974). On Lebesque-like extensions of finitely additive measures. Ann. Probability 2 456 - 463.

3. Dubins, Lester E. and Savage, Leonard J. (1976). Inequalities for stochastic processes: how to gamble if you must. Dover, New York.

4. Dubins, L.E. and Sudderth, W.D. (1977). Countably additive gambling and optimal stopping. Z. Wahrscheinlichkutstheorie und Verw. Gebiete 41 59 - 72 .

5. Dubins, L.E. and Sudderth, W.D. (1979). On stationary strategies for absolutely continuous houses. Ann. Probability 7 461 - 476 .

6. König, Denes (1936). Theorie der endlichen und undendlichen graphen. Akademische Verlagsgesellschaft, Leipzig.

7. Knuth, Donald E. (1973). The art of computer programming, volume 1. Addison-Wesley, Menlo Park, California.

8. Kuratowski, K. and Mostowski, A. (1976). Set theory. North Holland, Amsterdam.

9. Purves, Roger A. and Sudderth, William D. (1976). Some finitely additive probability. Ann. Probability 4 259 - 276.

10. Schäl, Manfred (1975). Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal. Z. Wahrscheinlichkeitstheorie verw. Gebiete 32 179 - 196.

11. Shreve, Steven E. and Bertsekas, Dimitri P. (1979). Universally measurable policies in dynamic programming. Math. of Operations Research 4 15 - 30.

12. Strauch, R.E. (1967). Measurable gambling houses. Trans. A.M.S. 126 64 - 72. (Correction, 130 184.)